

Prediction of air pollution peaks generated by urban transport networks

Margaret Bell¹, Angela S. Bergantino², Mario Catalano^{1*}, Fabio Galatioto¹

¹*Transport Operations Research Group (TORG), School of Civil Engineering and Geosciences, Newcastle University.*

²*Department of Economics, Management and Law (DISAG), University of Bari.*

Abstract

This paper illustrates the first results of an ongoing research for developing novel methods to analyse and simulate the relationship between transport-related air pollutant concentrations and easily accessible explanatory variables. The final scope of the analysis is to integrate the new models in traditional traffic management decision-support systems for a sustainable mobility of road vehicles in urban areas.

This first stage concerns the relationship between the mean hourly concentration of nitrogen dioxide and explanatory factors like traffic and weather conditions, with particular reference to the prediction of pollution peaks, defined as exceedances of normative concentration limits. Two modelling frameworks are explored: the Artificial Neural Network approach and the ARIMAX model. Furthermore, the benefit of a synergic use of both models for air quality forecasting is investigated.

The analysis of findings points out that the prediction of extreme pollutant concentrations is best performed by the integration of the two models into an ensemble. The neural network is outperformed by the ARIMAX model in foreseeing peaks, but gives a more realistic representation of the relationships between concentration and wind characteristics. So, it can be exploited to direct the ARIMAX model specification. At last, the study shows that the ability at forecasting exceedances of pollution regulative limits can be enhanced by requiring traffic management actions when the predicted concentration exceeds a threshold that is pretty high but lower than the normative one.

Keywords: air quality forecasting, exceedances of pollutant concentration limits, nitrogen dioxide, artificial neural network, ARIMAX model, ensemble techniques.

1 Research question and review of literature

It is broadly demonstrated that air pollution in urban areas is mainly due to the intense use of motorized transport for travelling, with particular regard to private cars and heavy goods vehicles. This is a top priority issue for transportation planners and public authorities, given the harmful effects of pollution to human health and the environment.

Numerous studies (Heinrich et al., 2005; Zhang et al., 2012) argue that acute exposure to air pollutants may cause serious temporary health concerns such as eye irritation, breathing difficulty, cardio-vascular problems, while chronic exposure may lead to damages to the body's immune, neurological, reproductive and respiratory systems, cancer and even premature death. In November 2014 the British Committee on the Medical Effects of Air Pollutants reported that air pollution may be responsible for as many as 60,000 early deaths in Britain each year. Also the environment is affected in terms of global climate change and adverse effects for plants and eco-systems (Seinfeld and Pandis, 2006; Zhang et al., 2012).

To protect human health and the environment, various national contexts throughout the world have issued guidelines and regulations. The United States Environmental Protection Agency (EPA) has set national ambient air quality standards for six pollutants: sulfur dioxide (SO₂), nitrogen dioxide (NO₂), carbon monoxide (CO), ozone (O₃), lead (Pb) and particulate matter (Seinfeld and Pandis, 2006).

* The contribution of Mario Catalano was carried out during his visiting in 2014 and 2015 at the Transport Operations Research Group of the University of Newcastle, School of Civil Engineering and Geosciences.

In Europe, over the last decades, the European Union has adopted an ample range of environmental measures to improve the quality of life for the Community's citizens. The final step of this legislative process is the Directive 2008/50/EC (EU, 2008), which has integrated an extensive body of laws establishing health-based concentration standards for a number of pollutants in outdoor ambient air. The European Commission has the task of ensuring that environmental law is applied by the Member States through infringement procedures.

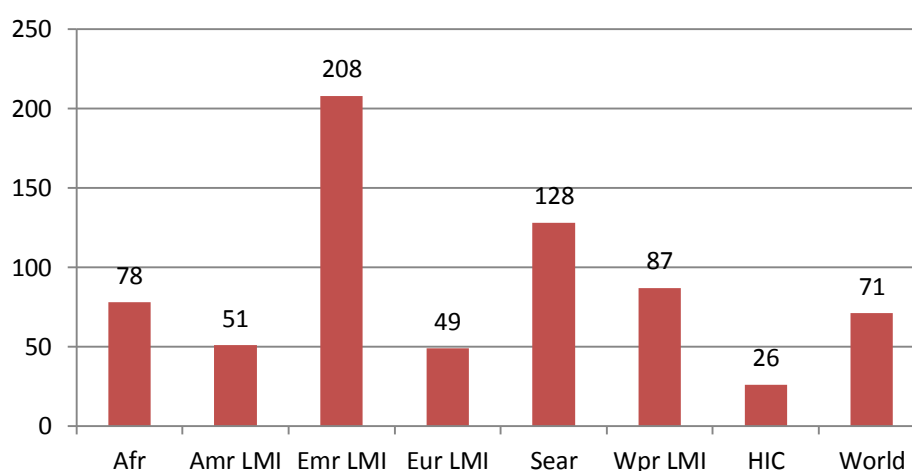
Long term measures like mode switch policies in favour of mass transit and public regulation on road use are pretty effective in abating atmospheric pollution in cities, but pollution peaks and the consequent exceedance of regulative concentration thresholds are often caused by substantial fluctuations of mobility patterns and weather conditions around their expected behaviours. Hence, air quality protection needs to be fine-tuned through the introduction in the local policy portfolio of further tools and actions to forecast extreme pollution events and manage traffic over short-term periods in order to prevent the predicted concentration peaks.

Given the above, this research has been started to investigate traffic-related air pollution modelling with the final aim of developing a real-time decision-support system for a more sustainable mobility of road vehicles in urban areas. In more detail, four main challenges will be addressed:

1. to develop a model to predict accurately the density of those airborne pollutants subject to normative standards for hourly or daily state of concentration, so as to permit local authorities to prevent their occurrence by real-time traffic management;
2. to explore the benefits in terms of accuracy and geographical transferability of models based on panel data, which are novel in the air quality modelling literature;
3. to build a method for predicting well in advance the yearly average of hourly mean concentrations, thus providing Local Authorities with a powerful tool to determine if and when to act in order to respect the environmental law, when this limits the behaviour of pollution over an annual period;
4. to experiment the synergic interplay between pollutant concentration forecasting and vehicular mobility microsimulation for an enhanced traffic management system based also on air quality targets.

By addressing the identified challenges, this research might be of strategic importance for many national contexts. To have an idea of the worldwide scale of atmospheric pollution problems, one could examine the 2014 version of the WHO (World Health Organization) Ambient Air Pollution database consisting mainly of urban air quality data, notably annual means of PM₁₀ and PM_{2.5}⁶ concentration for about 1,600 cities of 91 countries in the 2008-2013 period (WHO, 2014). As can be seen in Fig. 1, the world's annual mean levels of PM₁₀ by region range from 26 to 208 $\mu\text{g}/\text{m}^3$; the world's average is 71 $\mu\text{g}/\text{m}^3$ against the value of 20 $\mu\text{g}/\text{m}^3$ recommended by the WHO air quality guidelines (WHO, 2014). Particular concern is associated to the East side of the planet, where countries like China, India, Nepal, Bangladesh, Mongolia and, in the Mediterranean Area, Egypt, Iran, Jordan, Afghanistan, Pakistan far exceed the world's yearly mean density of PM₁₀.

PM₁₀ ($\mu\text{g}/\text{m}^3$)



⁶ Particles with diameter smaller than 10 and 2.5 microns, respectively.

Fig. 1. PM10 levels by region, for the last available year in the period 2008-2012. Amr: America, Afr: Africa; Emr: Eastern Mediterranean, Sear: South-East Asia, Wpr: Western Pacific; LMI: Low- and middle-income; HI: high-income.

The 2014 Air Quality in Europe Report (European Environment Agency, 2014) states that, in EU cities, exposure to atmospheric pollution levels exceeding the WHO air quality limits (in general stricter than the EU standards) is significantly high for various chemical agents. This over limit exposure regards 64% and 92% of the total EU-28 urban population in 2012 for PM10 and PM2.5, respectively. Moreover, in the case of ozone, in the same year, the exposure incidence rises to even 98% of people living in towns. There has been a clear decreasing trend, instead, for NO₂ concentration in many European countries over the last decade⁷, but, in the United Kingdom, the NO₂ levels have exceeded the relative WHO and EU target values persistently. This is confirmed by the fact that, in the early part of 2014, the European Commission launched legal proceedings against the UK for its failure to cut excessive levels of nitrogen dioxide (EU Press Release Database, 2014). Lastly, while exposure of the Europeans to CO concentrations above the EU and WHO thresholds is negligible, in the case of benzene (C₆H₆), around 10% of the EU-28 urban population is subject to pollution beyond the WHO levels and the percentage takes on the value of 37% in the case of SO₂.

This paper presents the early stage of our ongoing research, which refer to nitrogen dioxide, a toxic gas emitted by road vehicles, shipping, power generation, industry and households, which, even in the case of short term exposures (from 30 minutes to 24 hours), may cause adverse respiratory effects in healthy individuals. Furthermore, it is the main precursor for ground-level ozone, that is very harmful to human health. For NO₂, the EU environmental legislation sets two types of standard: the hourly mean concentration cannot go beyond the level of 200 µg/m³ more than 18 times each calendar year; the annual average of hourly concentrations is not allowed to exceed 40 µg/m³.

In particular, we have modelled the relationship between NO₂ hourly concentration and potential explanatory variables such as transport-related attributes, that influence emissions, and weather conditions, that are responsible for dispersion and transformation of pollutants.

As regards the scientific background of the research, few studies have appeared in the scientific literature on real time air quality forecasting near urban arteries, amongst which some are particularly interesting for this work, since they investigate the relationship between nitrogen oxides levels and meteorological and transport-related variables (Kukkon et al., 2003; Ming et al., 2009; Nagendra and Khare, 2006; Perez and Trier, 2001; Viotti et al., 2002). The leitmotiv of these studies is to consider the Neural Network, from the domain of Artificial Intelligence science, the most effective tool to predict air quality in urban areas. In some cases, this methodology is compared with other approaches, but they are usually linear regression models or deterministic models simulating the relevant physical processes. If also other pollutants are considered, it is possible to find air quality modelling works applying further statistical methods. For example, Arwa et al. (2014) tackle the issue of predicting particular matter concentration evaluating different models: multiple linear regression, quantile regression, generalised additive models and regression trees. Baur et al. (2004) compare the performance of quantile regression with multiple linear regression for predicting ozone concentrations. Kaushik and Melwani (2007) adopt the Seasonal Autoregressive Integrated Moving Average (ARIMA) model to forecast the daily levels of sulphur dioxide, nitrogen dioxide and suspended particulate matters.

Generally speaking, parametric and non-parametric statistical methods are more suitable for the description of complex relations between concentrations and potential predictors, and often present a higher accuracy, as compared to deterministic (physically-based) models, which are, furthermore, computationally expensive. However, statistical techniques are usually confined to the conditions occurring during the measurements and cannot be generalized to other areas with different chemical and meteorological characteristics. In addition, they fail to forecast concentrations during periods of unusual emissions and/or weather conditions that deviate significantly from the historical record (Zhang et al., 2012).

Given the above, in an attempt to develop an effective tool for predicting exceedances of NO₂ hourly concentration thresholds set by the EU, this research explores the statistical approach. For the particular case of nitrogen dioxide, a gap in the scientific literature has been identified in relation to the comparison between non-parametric statistical techniques, like the popular neural network, and sophisticated parametric methods as the Auto-Regressive Integrated Moving Averages with eXogenous inputs (ARIMAX) model (Hamilton,

⁷ Between 2003 and 2012, in EU-28, the ambient air NO₂ annual mean concentration dropped by 18% on the average. Only 8% of the EU-28 citizens live in areas where the annual WHO and EU thresholds for NO₂ were exceeded in 2012.

1994). Hence, the neural network and the ARIMAX framework have been compared with respect to NO₂ concentration forecasting through a set of indicators for missed exceedances and false alarms. In addition, based on successful experience in other fields⁸ (Bishop, 1995; Re and Valentini, 2012; VV. AA., 2008), the effectiveness of multimodel approaches have been evaluated. So, the forecasts from different models have been combined and the resulting impact on prediction accuracy has been quantified.

The remainder of the paper is structured as follows: section 2 describes the site where the data used in this work have been collected along with the dataset itself in terms of descriptive statistics and statistical properties of the involved time series; section 3 illustrates the theoretical foundations, specification and estimation of the models employed to forecast air quality; section 4 performs a comparative analysis of the prediction models based on statistical and categorical metrics; in the end, section 5 draws conclusions and points out new topics for future research.

2 Case study

This section gives details about the study area with respect to its geometric characteristics and the technologies used to collect air quality and traffic data. In addition, the dataset employed to derive insights into the research problem has been analysed with descriptive statistics and time series analysis techniques.

2.1 Air quality monitoring site

The study area chosen to perform the analysis is Marylebone Road in the City of London (see Fig. 2)⁹. Marylebone Road has three lanes each way, with the nearside lanes in both directions reserved for buses and taxis. The traffic along the corridor is controlled by a demand responsive signal control system termed SCOOT, Split Cycle Offset Optimization Technique (Hunt et al, 1981). The cabin for air quality measurement is located on the southern side of the road, on a spot where the road is characterised by a canyon H/W ratio of 0.86¹⁰ (see inset in Fig. 2). For Marylebone road, a reach dataset over a ten year period (1998-2007) is available, which contains traffic (flow and speed for each lane), weather and air quality data at an hourly resolution.

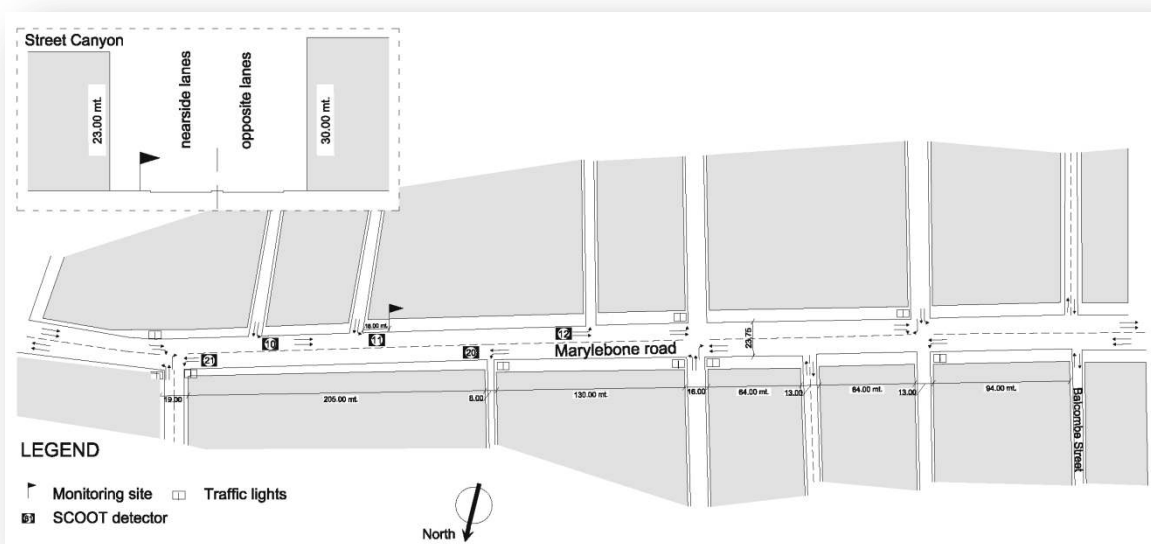


Fig. 2. The study area, Marylebone road in London, and the air quality monitoring site.

⁸ Such as machine learning science, astronomy, astrophysics, computer network intrusion detection, early diagnosis of diseases, face recognition.

⁹ For a preliminary analysis see: Bell et al. 2015.

¹⁰ The ratio of average buildings' height to road width (Tartaglia, 1999).

2.2 Descriptive analysis of the dataset

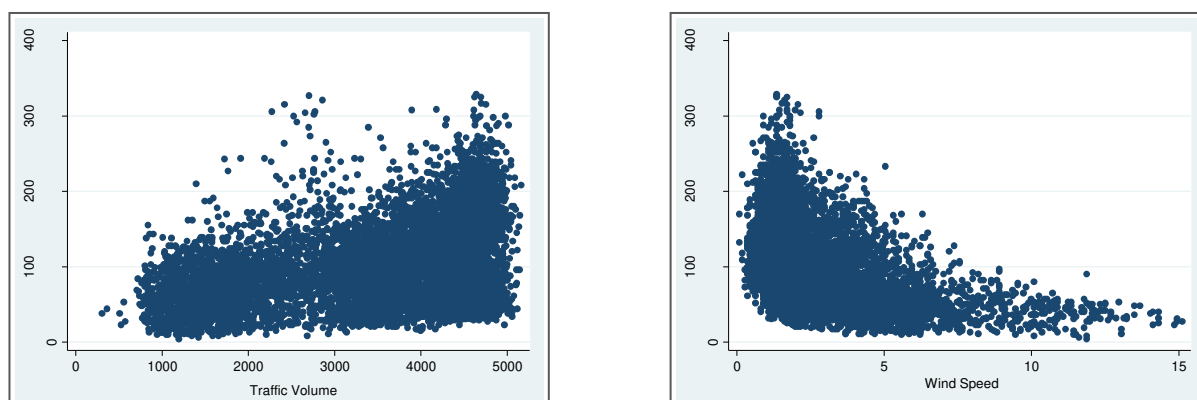
The analysis has concerned the relationship between the mean hourly concentration of NO₂ and explanatory factors like traffic and weather in a central spot of London, Marylebone Road, throughout the year 2007. Table 1 shows the main statistics describing the set of data collected in 2007 within the study-area. As can be seen in the table, the average of all hourly NO₂ concentrations is quite high (102.5); furthermore, it has been calculated that the 200 µg/m³ EU threshold was exceeded 457 times during 2007, against a maximum regulative limit of 18 times in a year. This makes the considered study-area a source of severe nitrogen dioxide air pollution, that needs long term policy-actions for sustainable transport, but also effective prediction models to support real-time traffic management. Such a problematic atmospheric condition may be probably ascribed to two main reasons: the high traffic volume which, from 7:00 to 21:00 o'clock, is greater than 4000 passenger car units/hour for 72% of time; the quasi-canyon layout of the monitored street, since the ratio of average buildings' height to road width is around 0.9 (Tartaglia, 1999).

Table 1. Descriptive statistics for the set of air quality, transport and weather data collected in 2007 within the study-area.

| Variables | Observations | Average | Min | Max | Standard Deviation |
|---|--------------|---------|-------|--------|--------------------|
| Hourly mean concentration of NO ₂ (µg/m ³) | 8584 | 102.5 | 4 | 329 | 53.1 |
| Total traffic volume* (passenger car units/hour) | 8734 | 3505.7 | 303.7 | 5161.5 | 1173.2 |
| Hourly mean speed of wind (km/h) | 7993 | 2.8 | 0.1 | 15.1 | 2.1 |
| Hourly mean direction of wind (North degrees) | 7993 | 203.9 | 2.9 | 358.1 | 82.4 |
| Hourly mean temperature (centigrade degrees) | 7993 | 13.5 | -0.8 | 29.7 | 5.2 |

* The original data on traffic were disaggregated by six vehicle classes: Motorcycle, Car or Light Van (length < 5.2m), Car and trailer, Rigid Lorry, Heavy Van (length ≥ 5.2m) or mini-bus, Articulated Lorry, Bus or Coach. The measured flows for each category have been turned into passenger car equivalents through the following multiplicative parameters (Lavecchia et al., 2007): 1 for motorcycles and cars, 1.5 for light vehicles (5m < length < 7.5m), 2 for heavy vehicles (7.5m < length < 12.5m), 3 for very heavy vehicles (length > 12.5m).

Fig. 3 shows through scatter graphs the relationships between the hour-based mean density (µg/m³) of NO₂ at Marylebone road in 2007 and the explanatory variables taken into account, which are the hourly averages of traffic volume (passenger car units/hour), wind speed (km/h), wind direction (North degrees) and temperature (centigrade degrees). As can be observed, traffic flow has a positive impact on concentration, while wind speed is inversely related to pollutant density; moreover, the relationship between pollution and wind direction looks quite complex and needs further investigation, whereas temperature seems negligible. Especially in the case of weather variables, the influence on concentration follows a non-linear behaviour.



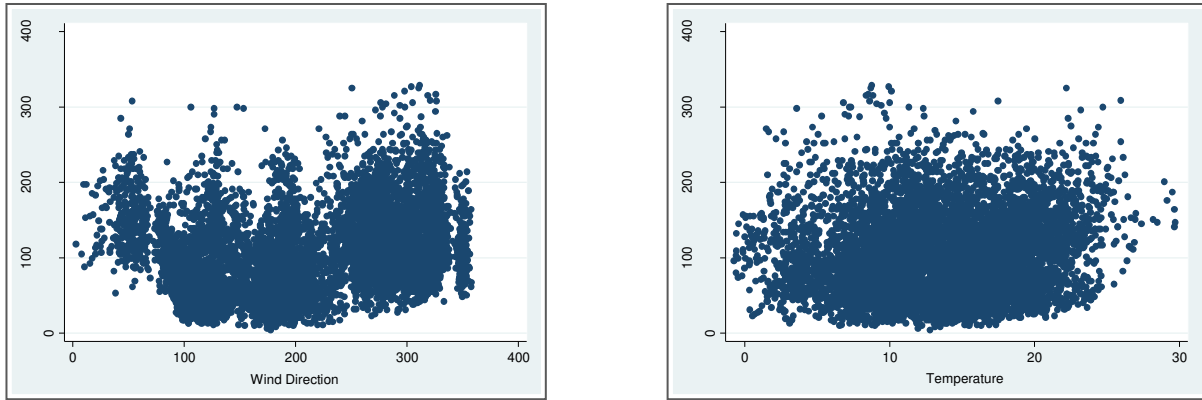
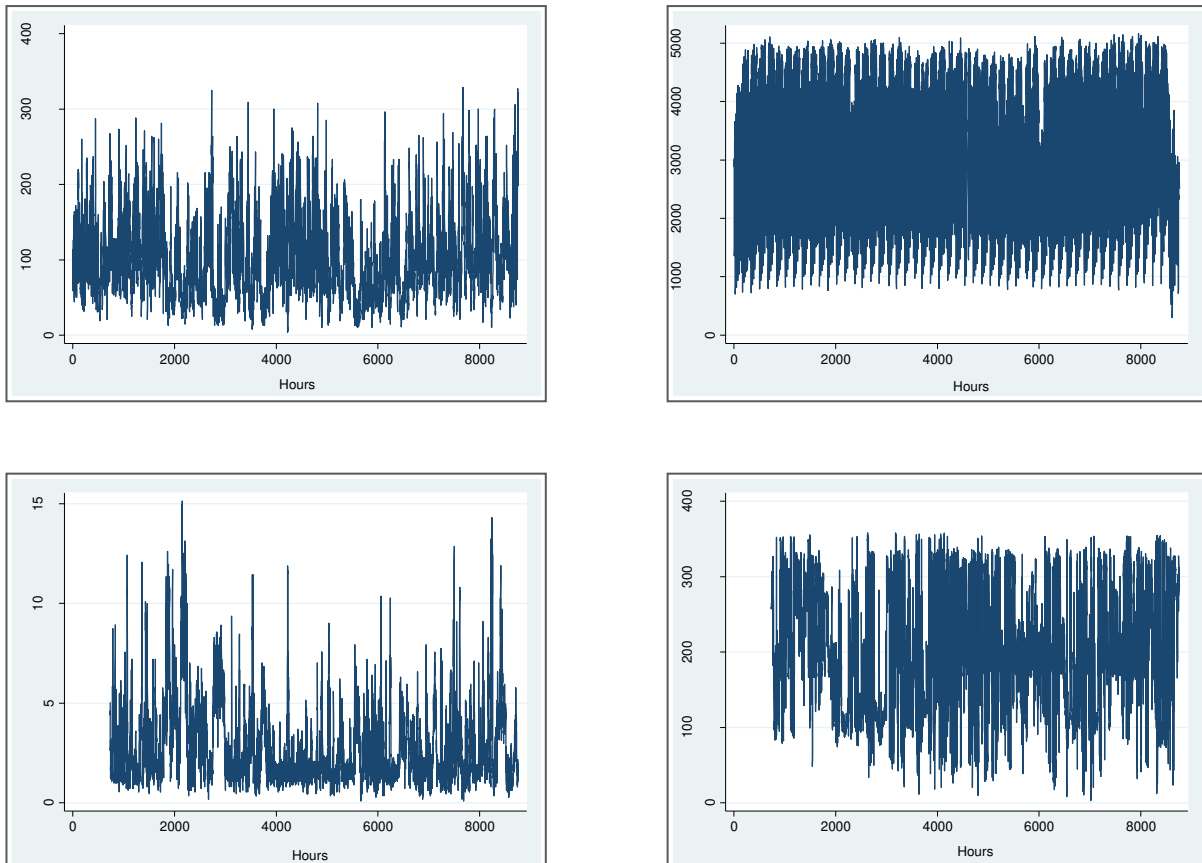


Fig. 3. Scatter plots depicting the relationships between NO₂ mean concentration and its explanatory variables for the set of hourly observations recorded in 2007 at Marylebone road.

2.3 Time series analysis

The time series of the considered variables has been analysed to check their stationarity, a statistical property of the series-generating stochastic process which requires its mean and autocovariance to be date-independent. It is also a regularity condition which enables the development a model to explain a certain variable with its past values and a set of exogenous predictors (Hamilton, 1994). We have checked whether this property holds in our case-study through graphical inspections and formal tests.

As Fig. 4 displays, the time mean and variance of the variables of interest look quite stable through the year, except for the hourly averages of vehicle speed and temperature. In the first case, there is a turning point in the time series behaviour in the early part of the year. Temperature, as expected, shows an increasing trend in the first half of the year and a decreasing one in the second half.



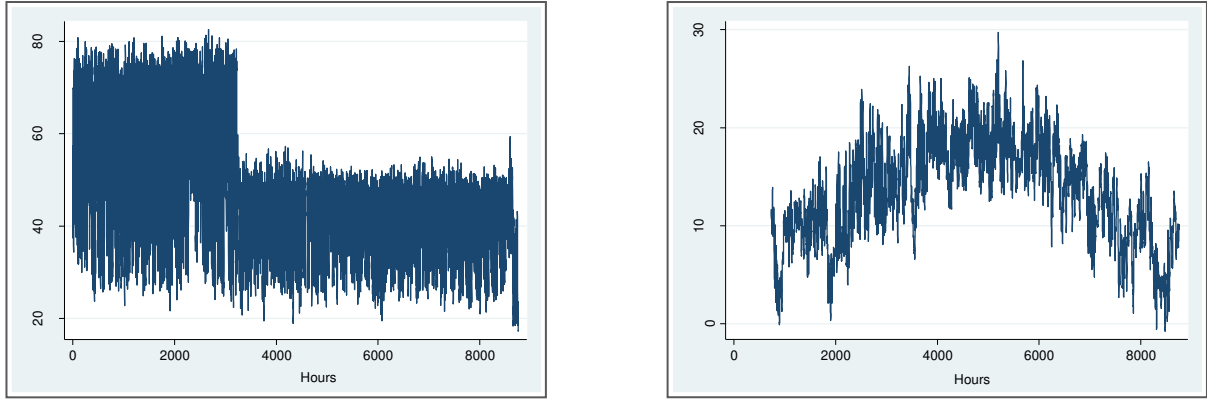


Fig. 4. Line plots depicting the hourly averages of NO₂ density (µg/m³), traffic volume (passenger car units/hour), wind speed (km/h), wind direction (North degrees), vehicle average speed (km/h) and temperature (centigrade degrees), in 2007 at Marylebone road.

Speed and temperature as non-stationary inputs cannot be employed to explain NO₂ hourly density, which is instead stationary. The stationary nature of NO₂ concentration, traffic volume, wind speed and direction, which is pretty clear in the respective line plots, is confirmed by the Augmented Dickey-Fuller (ADF) unit root test (Hamilton, 1994). On one hand, under the most general null hypothesis, this test assumes that the model representing the true behaviour of a given time series y_t can be formulated as a random walk with drift:

$$y_t = \alpha + \rho(= 1) \cdot y_{t-1} + \varepsilon_t \quad (1)$$

where,

y_t : dependent variable (in our case, the hourly mean concentration of NO₂), assumed to depend upon its first order lagged value y_{t-1} by coefficient ρ equal to one (unit root), which implies non stationarity in variance;
 α : drift term which makes y_t expressable as the sum of a linear time trend and a series of random impulses ($y_t = \alpha \cdot t + \sum_{i=0}^t \varepsilon_i$);
 ε_t : independently and identically distributed error terms with zero mean (white noise);

On the other hand, the most general alternative hypothesis of the test is the stationary behaviour of the series around a deterministic linear time trend:

$$y_t = \mu + \delta \cdot t + \rho(< 1) \cdot y_{t-1} + \varepsilon_t \quad (2)$$

where

$(\mu + \delta \cdot t)$ is the trend;
 ρ is smaller than one, thus making the fluctuations of the series around its trend be the manifestation of a first-order autoregressive model;

Since model (2) might not capture fully the underlying serial correlation (the autocorrelation could go far beyond the first order level), the ADF test fits through the Ordinary Least Square technique a transformation of model (2):

$$\Delta y_t = \mu + \delta \cdot t + \beta \cdot y_{t-1} + \gamma_1 \cdot \Delta y_{t-1} + \dots + \gamma_k \cdot \Delta y_{t-k} + \varepsilon_t \quad (3)$$

where,

$\Delta y_t = y_t - y_{t-1}$: first difference of variable y_t ;
 $\gamma_j \cdot \Delta y_{t-j} = \gamma_j \cdot (y_{t-j} - y_{t-j-1})$: additional elements to capture the likely presence of further serial correlation in y_t ;

Testing $\beta = 0$ amounts to testing $\rho = 1$ (Hamilton, 1994) or, equivalently, that y_t follows a unit root process, which means that it is not stationary and, hence, to become stationary, it has to be differenced subtracting from each element of the series its first-order lagged value. The test has been performed for the hourly averages of NO₂ density, traffic volume, wind speed, wind direction. Several values of lag k have been experimented and in each case the null hypothesis ($\beta = 0$) has been rejected, thus confirming the output of the preliminary graphical inspection of the above variables. Table 2 presents the test results for lag $k = 24$, which can take account of possible forms of autocorrelation till a day before the hour of interest. Since the series under consideration do not show a trend, but have a nonzero mean, under the null hypothesis, the value of the drift (α) has been set at zero, while, under the alternative hypothesis, only the δt term of regression (3) has been dropped. Table 2 shows that, in each instance, the test statistic is far smaller than the critical values at 1%, 5% and 10% significance levels, which makes the assumption of stationarity highly plausible.

Table 2. Augmented Dickey-Fuller unit root test results if lag $k = 24$.

H₀: $\beta = 0$ and $\alpha = 0$;

H₁: $\delta = 0$

| Variables | Test Statistic | Critical Value | Critical Value | Critical Value |
|---|----------------|----------------|----------------|----------------|
| | | 1% | 5% | 10% |
| Hourly mean concentration of NO ₂ ($\mu\text{g}/\text{m}^3$) | -9.090 | -3.430 | -2.860 | -2.570 |
| Total traffic volume (passenger car units/hour) | -8.905 | -3.430 | -2.860 | -2.570 |
| Hourly mean speed of wind (km/h) | -9.143 | -3.430 | -2.860 | -2.570 |
| Hourly mean direction of wind (North degrees) | -9.075 | -3.430 | -2.860 | -2.570 |

To specificate a model for the time series of NO₂ hourly concentration, it is useful to study also its autocorrelation. This has been done through the visual analysis of the total and partial¹¹ autocorrelation functions. In Fig. 5, the behaviour of total autocorrelation reveals a weak seasonal effect: in fact, as the lag extends, the positive value of autocorrelation declines, but with a reversal occurring whenever the time interval gets to a multiple of twentyfour hours. In addition, the partial autocorrelation highlights a strong dependence of concentration upon its value one hour back, whereas the other direct relationships with past concentrations are negligible. The seasonal effect can be interpreted considering that road mobility follows a cyclical pattern that stems from people's routine and, hence, is incorporated into the behaviour over time of all the pollutants strictly related to transport such as NO₂. The positive correlation with the mean density of the previous hour, instead, could be reasoned out considering that the amount of pollutant emitted during an hour requires a certain time to be dissipated, so a fraction of it represents the background for the next hour. Moreover, the background pollution in an hour contains not only what has been emitted and not dispersed during the preceding hours, but also what is emitted in the hour of interest by sources different from transport, mainly residences in our study-area. The impact of changes in this part of the background on concentration is likewise taken into account by the average density of the previous hour. In fact, variations in the emission pattern of residential activities with respect to its mean behaviour can be considered persistent at least within the same day.

¹¹ The partial autocorrelation at lag k measures the correlation between x_t and x_{t+k} after the effects of $x_{t+1}, \dots, x_{t+k-1}$ have been removed.

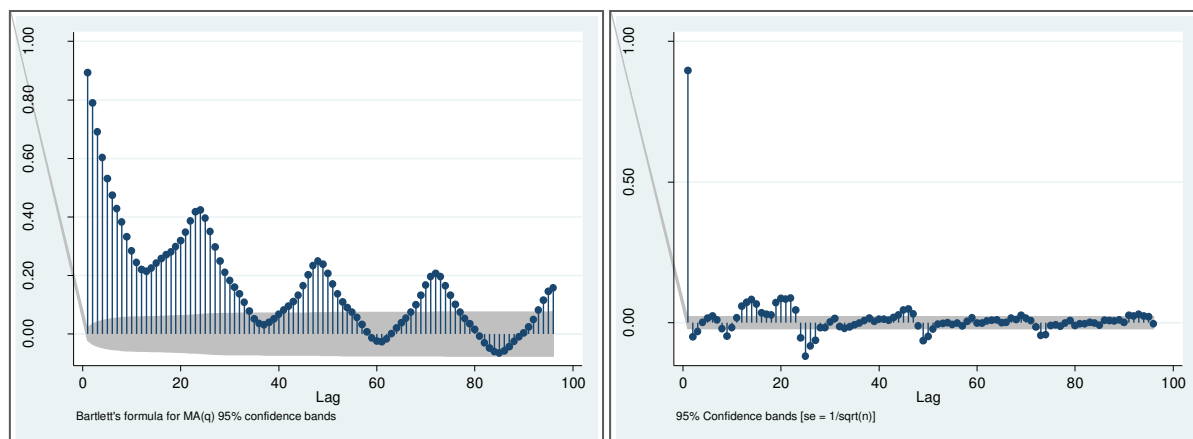


Fig. 5. Graphes of the total and partial autocorrelations for hourly average of NO₂ density ($\mu\text{g}/\text{m}^3$) in 2007 at Marylebone road.

3 Modelling approaches explored

As stated previously, this work compares the performance in forecasting NO₂ concentration of two families of models: the Artificial Neural Networks (ANN) and the Auto-Regressive Integrated Moving Averages models with eXogenous variables (ARIMAX). Both belong to the statistics domain, which uses mathematical functions built on empirical data to predict concentration. Unlike the ARIMAX methodology, the ANN approach does not require specific assumptions on the mathematical relationships between the pollutant density and its explanatory factors. There is very little literature on this type of comparative analysis, especially with reference to the prediction of extreme levels of pollution.

The comparison is based on the dataset of NO₂ hourly mean concentrations ($\mu\text{g}/\text{m}^3$) measured throughout the year 2006 at Marylebone road in London. The variables to explain the state of concentration are representative of weather and traffic conditions during the hour of interest: traffic volume (passenger car units/hour), hourly mean speed of wind (km/h), hourly mean direction of wind (North degrees). Besides, Either models embody the double form of autocorrelation emerging from the analysis of correlograms (Fig. 4), that is the dependency of concentration upon its past values one hour and one day back.

So, to use these models for one step ahead concentration forecasting, traffic flow and wind characteristics have to be foreseen on an hourly basis. For the former, assignment of historical hourly O/D matrices improved with real-time traffic counts to the future state of road network can be performed. For meteorological factors, very precise short-term predictions can be obtained by the providers of weather forecasts that regularly process the meteorological data collected throughout the country.

The following subsections illustrate the concepts underpinning the two approaches in comparison and the frameworks of the specific models derived from the data.

3.1 Neural network

The neural network is an artificial intelligence-based technique that mimics the human brain behaviour in a learning process. The important feature of an ANN is its adaptive nature: "learning by examples" is the approach used by this method to accomplish classification and regression tasks. This makes neural networks virtually applicable in every situation in which some of the relationships between a response variable and its predictors are very complex and cannot be easily outlined based on a priori knowledge and theoretical considerations.

The Multilayer Perceptron (MLP) architecture has proved to be the most suitable class of neural networks for air quality forecasting in previous studies (Nagendra and Khare, 2002). It consists of a system of layered and interconnected nodes or neurons. They form an input layer, one or more hidden layers and an output layer with nodes in each layer connected to all nodes in neighbouring layers (Bishop, 1995). The input layer neurons provide input signals to the hidden layer, where each node sums the inputs, processes the result with a nonlinear transfer or activation function (logistic or hyperbolic tangent) and then distributes it to the output layer. The output layer computes the dependent variable value in a similar manner. Each neuron-to-neuron connection is associated to a specific weight. MLP has the ability to learn through training, which requires a

series of input vectors and associated outputs. During training, the output from the MLP is compared with the desired value, an error signal is propagated back through the network and the magnitude of this error is used to adjust the weights iteratively till a stopping criterion is met. Besides a set of observations for training (training set), a further set of new data (validation or selection set) is used to limit the effective network complexity in favour of its generalization performance, which is the prediction or classification accuracy with respect to unknown cases. In fact, during a typical network learning session, the training set error generally decreases as a function of the number of iterations. On the contrary, the error measured with respect to independent data (validation set) often shows a decrease at first, followed by an increase as the network starts to over-fit. So, halting training at the point of smallest error with reference to new data should result in achieving the best network generalization performance.

It is common practice in the application of neural networks to train many different candidate networks and then to select the best on the basis of validation set performance. However, a selection based only on the validation set could be misleading, because the network performance has a random component due to data noise. Hence, the choice of a specific network structure should be confirmed by its performance in relation to a third independent set of data (test set).

Fig. 6 depicts the neural architecture that has been selected to simulate the hourly average of NO_2 density ($\mu\text{g}/\text{m}^3$) in 2007 at Marylebone road. The independent variables are the mean concentrations one hour and one day back along with the non-lagged hourly averages of traffic volume (passenger car units/hour), wind speed (km/hour) and wind direction (North degrees). The neural network design has been performed through a number of experiments with different architectures (three and four layer MLP networks) and the network with the best selection set performance has been chosen. At each learning round, to make the candidate networks as diverse as possible, resampling has been used to form the training, selection and test subsets from the whole sample of data. In more detail, the Monte Carlo technique has been adopted drawing at random the training and selection subsets, while fixing the test subset once and for all. To decide the number of hidden units, weight regularization (Bishop, 1995) has been performed adding an extra term to the error function for penalizing and reducing weights giving a small contribution to the network performance. This procedure prunes entire hidden units when their fanning-out weights are below a fixed threshold, thus limiting the curvature of sigmoid activation functions and improving the network generalization skill.

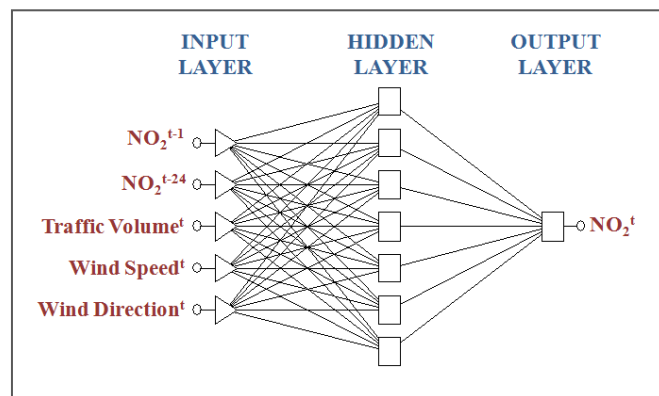


Fig. 6. Neural Architecture selected to forecast the hourly average of NO_2 density ($\mu\text{g}/\text{m}^3$) in 2007 at Marylebone road.

Table 3 reports the Pearson-R correlation coefficient between the model's predictions and the observed values and the performance of the best network for each subset of data. The high correlation reveals the ability of the neural model at reproducing the historical behaviour of concentration accurately. Every performance indicator is computed as the ratio of the prediction error standard deviation to the standard deviation of the specific subsample of data. A ratio remarkably below 1 indicates that the network has performed far better than a simple mean estimator and its use in forecasting is then justified. As Table 3 exhibits, for the chosen network the above ratio is quite low in relation to each part of the databank, which signals a good level of performance. In addition, the selection performance is better than that for the training set, which means that the network has not overlearned.

Table 3. Performance of the best MLP neural network.

| Observations | Train Performance | Selection Performance | Test Performance | Obs.-Pred. Correlation |
|--------------|-------------------|-----------------------|------------------|------------------------|
| 7563 | 0.403 | 0.365 | 0.400 | 0.920 |

Fig. 7 displays the response surfaces plotted in relation to two couples of input variables. To construct every response surface, the two chosen inputs vary from their observed minima to their observed maxima, while all other inputs are held at a fixed value (their averages). In one case, the picture visualizes the dependency of the NO₂ hourly density upon its lagged values showing a strong positive impact of the mean concentration in the preceding hour and a slight effect of the corresponding hourly concentration of the previous day.

In the other case, it is shown that the relationship between concentration and wind characteristics is clearly non-linear: as the wind speed rises, concentration reduces at a decreasing marginal rate, whereas the dependency upon the wind direction is represented by a U-shaped curve.

The second relation, in particular, can be ascribed to the canyon layout of Marylebone road. The part of this artery just off the monitoring device presents a H/W ratio¹² close to 1, which indicates the possibility of a canyon-street behaviour of transport-related pollution (Tartaglia, 1999). This means that, when the wind speed is at least 1 m/sec and its direction forms with the road normal axis an angle in the -45°-45° range, a vortex is generated inside the canyon-road and the pollutant concentration on the windward side is far higher than that on the leeward side. Marylebone road has a Southwest-Northeast alignment and the pollution receptors are located on its southern boundary (see Fig. 2). It follows that, when the wind comes from South, the set of sensors is on the windward side, so one should observe an increasing pollution concentration as the wind direction gets closer to 360 (or zero) North degrees. On the contrary, if the wind comes from South, the sensors are on the leeward side, hence concentration should be decreasing as the wind gets closer to 180 North degrees. The trained neural network is able to simulate this complex behaviour of pollution, which demonstrates its good level of realism.

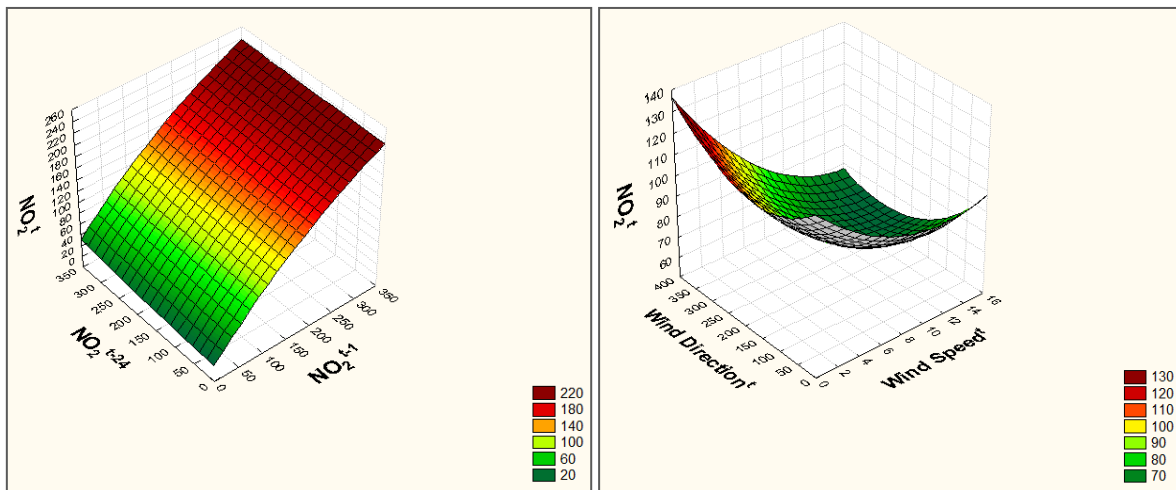


Fig. 7. Response surfaces describing how the trained MLP neural network simulates the dependency of the NO₂ hourly density ($\mu\text{g}/\text{m}^3$) upon, in one case, its lagged values (one hour and one day back) and, in the other case, the wind characteristics (speed in km/hour and direction in North degrees).

In the end, Fig. 8 points out that the traffic volume, as expected, has a positive impact on concentration according to a mathematical relationship that slightly deviates from the linear function.

¹² The ratio of average buildings' height to road width (Tartaglia, 1999).

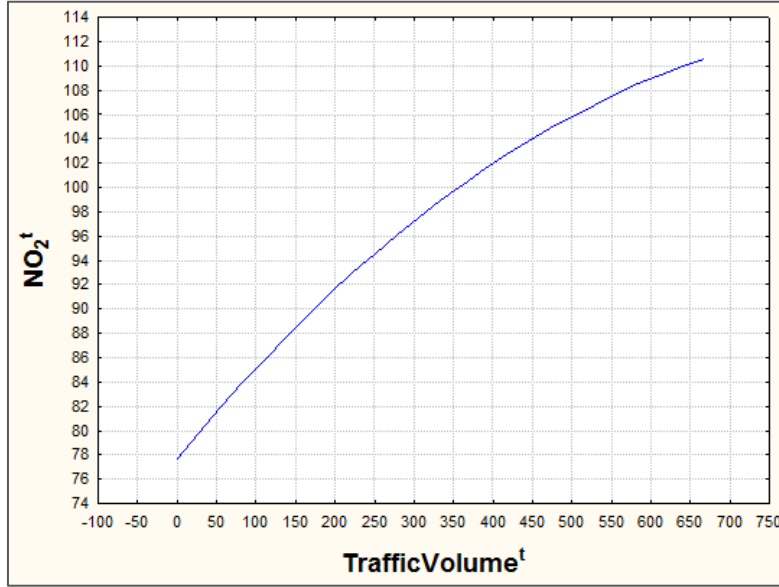


Fig. 8. Response line describing how the trained MLP neural network simulates the dependency of the NO₂ hourly density (µg/m³) upon the traffic volume (passenger car units/hour).

3.2 Statistical model

For the prediction of NO₂ hourly concentration at Marylebone road, also an ARIMAX model has been developed. This represents the underlying data-generating stochastic process as the integration of two main components: one captures the relationship between the dependent variable and its past manifestations (autoregressive part), the other incorporates the effect on the dependent variable of a finite series of random impulses (moving average part). Formally, an ARIMAX model, with p autoregressive (AR) elements and q moving average (MA) terms, can be written as follows:

$$(1 - \rho_1 L^1 - \rho_2 L^2 - \dots - \rho_p L^p) \cdot (y_t - \mathbf{x}'_t \cdot \boldsymbol{\beta}) = (1 + \theta_1 L^1 + \theta_2 L^2 + \dots + \theta_q L^q) \cdot \epsilon_t \quad (4)$$

or, more succinctly:

$$\rho(L^p)(y_t - \mathbf{x}'_t \cdot \boldsymbol{\beta}) = \theta(L^q)\epsilon_t \quad (5)$$

where

ρ_j : coefficient of the j^{th} -order AR element;

θ_j : coefficient of the j^{th} -order MA element;

L^j : lag operator transforming a variable at time t into its j^{th} -order lagged manifestation ($L^j z_t = z_{t-j}$);

y_t : dependent variable at time t ;

\mathbf{x}'_t : vector of exogenous covariates at time t ;

$\boldsymbol{\beta}$: vector of coefficients;

$\epsilon_t \sim N(0, \sigma^2)$, meaning that it is a white noise disturbance;

$\rho(L^p) = (1 - \rho_1 L^1 - \rho_2 L^2 - \dots - \rho_p L^p)$;

$\theta(L^q) = (1 + \theta_1 L^1 + \theta_2 L^2 + \dots + \theta_q L^q)$.

If the considered series shows a cyclical behaviour, model (5) can be turned into a seasonal ARIMAX or SARIMAX model by the introduction of multiplicative terms both for the AR part and for the MA component:

$$\rho(L^p)\rho_s(L^p)(y_t - \mathbf{x}'_t \cdot \boldsymbol{\beta}) = \theta(L^q)\theta_s(L^q)\epsilon_t \quad (6)$$

where the seasonal symbols (those with subscript s) have the same meaning as their non-seasonal counterparts, but apply to the series every s periods.

The SARIMAX model for the 2007 time series of NO₂ concentrations at Marylebone road has one AR term and one MA element both for the non-seasonal part and for the seasonal one, the latter being based on a periodicity of 24 hours (recall subsection 2.3). The exogenous regressors are again the traffic volume (autos/hour), the wind speed (km/hour) and the wind direction (North degrees). Given the above, formula (6) can be particularized as follows:

$$\rho(L^1)\rho_{24}(L^1)(y_t - \mathbf{x}'_t \cdot \boldsymbol{\beta}) = \theta(L^1)\theta_{24}(L^1)\epsilon_t \quad (7)$$

Table 4 exhibits the results of the SARIMAX model estimation, conducted by the maximum likelihood technique and using the Kalman filter via the prediction error decomposition (Hamilton, 1994). This approach handles missing data (for the dependent variable and/or its covariates) by continuing the state-updating recursions of the Kalman filter even if the contribution to estimation from the sample is partial or even null. All the estimated parameters have correct signs and high statistical significance except the constant.

The main limitation of the SARIMAX model is its inability to reproduce the non-linear relationship between concentration and the wind characteristics, in particular with regard to direction. In this case, in fact, a linear positive impact on pollution is very far from the more plausible parabolic behaviour found out with the neural network and explained as a canyon street effect.

In line with the Box and Jenkins' guidelines (Box et al., 2008), the SARIMAX residuals have been analysed to check whether they can be reasonably considered gaussian and independently distributed as required by the theory. Fig. 9 depicts the frequency distribution of residuals along with their autocorrelation function for the first 24 lags. As can be observed, the assumption of a white noise behaviour of residuals is supported by the empirical evidence, which validates the SARIMAX model.

Table 4. The SARIMAX model for the 2007 series of NO₂ hourly concentrations at Marylebone road.

| Hourly Averages of regressors | Coefficient | Standard Error | z | P> z | [Conf. Int. 95%] | |
|--|-------------|----------------|---------|-------|------------------|--------|
| Traffic Volume (passenger car units/h) | 0.074 | 0.007 | 10.530 | 0.000 | 0.061 | 0.088 |
| Wind Speed (km/h) | -8.805 | 0.348 | -25.330 | 0.000 | -9.486 | -8.123 |
| Wind Direction (North degrees) | 0.025 | 0.005 | 4.990 | 0.000 | 0.015 | 0.035 |
| Costant | -0.118 | 0.123 | -0.960 | 0.336 | -0.358 | 0.122 |
| NO ₂ Concentration 1 Hour Back (µg/m ³) | 0.873 | 0.009 | 98.050 | 0.000 | 0.856 | 0.890 |
| NO ₂ Concentration 1 Day Back (µg/m ³) | 0.074 | 0.016 | 4.670 | 0.000 | 0.043 | 0.104 |
| Moving Average Component | -0.098 | 0.020 | -4.810 | 0.000 | -0.138 | -0.058 |
| Seasonal Moving Average Component | -0.950 | 0.007 | 130.530 | 0.000 | -0.965 | -0.936 |
| Error Standard Deviation | 20.378 | 0.288 | 70.830 | 0.000 | 19.814 | 20.942 |

LN (Likelihood) = -33783.07
 N. observations = 7563
 Wald X²(7) = 31839.73

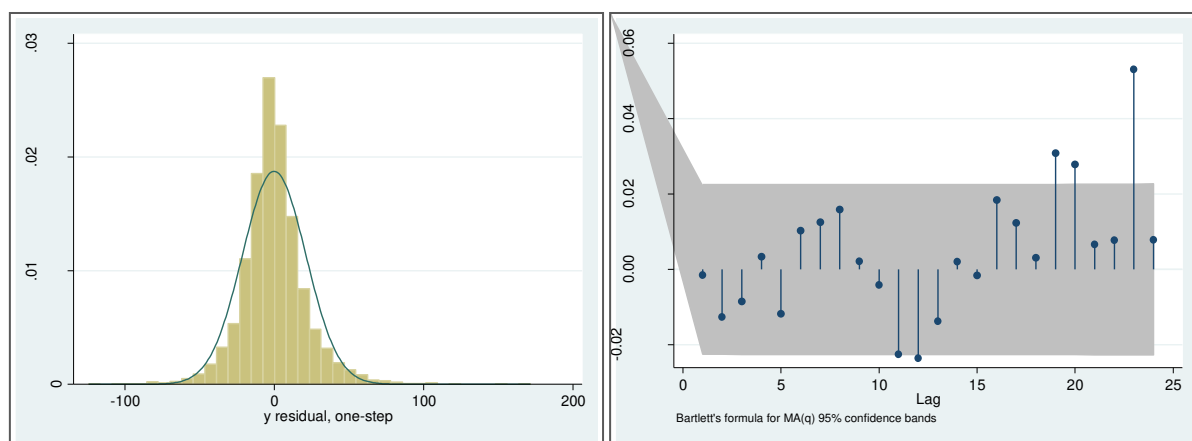


Fig. 9. Analysis of the SARIMAX model's residuals: distribution and autocorrelation.

4 Comparative analysis

This section compares the two forecasting models in relation to a sample of observations that has not been involved in model building. The comparison is based on various criteria, evaluated through statistical and categorical metrics: the extent to which observed and simulated concentrations covary, measured by the Pearson-R correlation coefficient; the global accuracy, quantified by the mean absolute percent error (MAPE); the ability to predict exceedances of fixed threshold densities, measured by the percentage of the actual exceedances forecasted by the models and the false alarm ratio, that is the ratio of the predicted exceedances that actually did not occur to all foreseen exceedances. Moreover, to explore the potential advantage of ensemble forecasting, the predictions of the two models have been combined and the resulting performance has been evaluated.

4.1 The dataset for comparison

The dataset used to compare the estimated models refers to the year 2006. For this period, the annual average of NO₂ hourly mean concentrations is very close to that calculated for the year 2007 (110.6 versus 102.5), but the number of cases in which the 200 µg/m³ density was exceeded is 1.5 times greater (686 versus 457). As emerges from Table 5, the base statistics describing the 2006 set of observations do not differ a great deal with respect to the 2007 one (Table 1).

Table 5. Descriptive statistics for the set of air quality, transport and weather data collected in 2006 within the study-area.

| Variables | Observations | Average | Min | Max | Standard Deviation |
|---|--------------|---------|-------|--------|--------------------|
| Hourly mean concentration of NO ₂ (µg/m ³) | 8518 | 110.6 | 2 | 403 | 56.6 |
| Total traffic volume (passenger car units/hour) | 5637 | 3377.2 | 721.7 | 4899.2 | 1158.1 |
| Hourly mean speed of wind (km/h) | 6871 | 3.0 | 0.2 | 17.6 | 2.4 |
| Hourly mean direction of wind (North degrees) | 6871 | 217.9 | 1.0 | 358.7 | 85.8 |
| Hourly mean temperature (centigrade degrees) | 6865 | 13.9 | -2.1 | 37.3 | 7.1 |

4.2 Evaluation of model forecasting performance and sensitivity analysis

As already emphasized, the artificial neural network and the SARIMAX model have been compared with respect to the ability to predict the NO₂ concentrations recorded at Marylebone road in 2006. In order to provide initial and tentative conclusions on the potential benefit of ensemble forecasting, the two models have been integrated into a mixture model by taking for each case the maximum of the concentrations predicted by both of them. The choice of the *max* functional form for the ensemble is motivated by the two

models' tendency to underestimate pollution peaks. In fact, for the instances in which the threshold density ($200 \mu\text{g}/\text{m}^3$) is exceeded on average during an hour, the mean ratio of the predicted concentration to the actual one is 0.89 in the SARIMAX case and 0.86 for the neural network; it becomes a bit higher (0.92) if the two models are merged into an ensemble.

Fig. 10 visually outlines the distributions of measured and simulated concentrations. The graphic tool employed is the box plot, which represents the 25th and 75th percentiles with the the bottom and top lines of the box, respectively. Moreover, the line intersecting the box identifies the median of the distribution, while the two whiskers provide a measure of the data's spread. The isolated points out of the spread represent outliers. As can be seen in the picture, all the models reproduce the observed yearly distribution of concentrations quite well. In addition, it is clearly displayed by the upper whiskers that all models underestimate the high level of concentrations, even though the ensemble mitigates this problem.

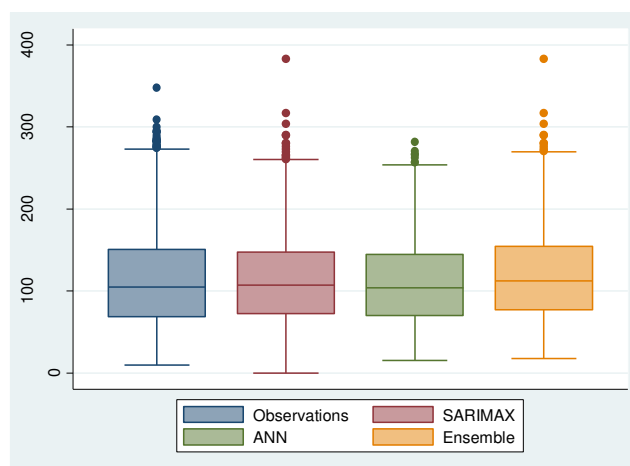


Fig. 10. Box Plot of NO_2 concentration distribution in 2006 at Marylebone road: observations versus simulations.

Table 6 compares all models with five performance indicators. First, two statistical indices are used to measure the correlation and the mean absolute percent difference between observed and estimated densities. With reference to these indicators, the three considered models do not differ a great deal and can be deemed fairly effective.

Second, two categorical indices are used to evaluate the model ability to foresee the exceedance of the $200 \mu\text{g}/\text{m}^3$ limit. One indicator is the percentage of the real exceedances forecasted by the models. From this point of view, the ensemble approach yields a rather better performance, since the *max* operator lowers the tendency to underestimate extreme values of pollution. The other indicator is the share of the predicted exceedances (alarms) that did not occur (false alarm ratio). All models present good levels of this performance measure with a percentage of false alarms (predicted exceedances that did not occur) varying in the 30-37% range. Furthermore, when a false alarm is generated by any model, the corresponding real concentration is quite high on average (around $180 \mu\text{g}/\text{m}^3$), which makes such an alarm useful anyway.

As the models underestimate pollution peaks, their ability to foresee exceedances could be enhanced by establishing an alarm threshold smaller than $200 \mu\text{g}/\text{m}^3$, but high anyway, say $180 \mu\text{g}/\text{m}^3$. This means planning to intervene through traffic management whenever a model forecasts for the next hour a density beyond $180 \mu\text{g}/\text{m}^3$. By doing so, one could expect better values for the percentage of the actual exceedances captured by the models, since the real concentrations slightly above $200 \mu\text{g}/\text{m}^3$ and estimated by the models at lower values but higher than $180 \mu\text{g}/\text{m}^3$ would trigger traffic management actions under the new approach. Table 7 confirms this assumption, showing that models' performance in foreseeing historical exceedances improve by even 22 percent points at the expense of an acceptable rise in the incidence of false alarms (from 30-37% to 45-51%). Furthermore, the mean density when a false alarm is issued is still pretty high (170 - $173 \mu\text{g}/\text{m}^3$), which guarantees an useful traffic management intervention.

In terms of actual exceedances foreseen by the models, better results could be obtained with lower alarm thresholds to the detriment of the false alarm ratio; at any rate, $180 \mu\text{g}/\text{m}^3$ leads to a satisfactory balance between the two conflicting evaluation criteria.

Table 6. Evaluation of forecasting performance if the alarm threshold is $200 \mu\text{g}/\text{m}^3$.

| Modelling Approach | Correlation | MAPE (%) | Predicted Exceedances (%) | False Alarm Ratio (%) | Mean Density for False Alarms ($\mu\text{g}/\text{m}^3$) |
|---|-------------|----------|---------------------------|-----------------------|--|
| SARIMAX model | 0.91 | 18.62 | 54.61 | 33.86 | 177.85 |
| Artificial Neural Network | 0.91 | 16.53 | 49.38 | 30.40 | 180.59 |
| Esemble: max [Pred _{SARIMAX} , Pred _{ANN}] | 0.92 | 19.32 | 61.95 | 36.77 | 177.66 |

Table 7. Evaluation of forecasting performance if the alarm threshold is $180 \mu\text{g}/\text{m}^3$.

| Modelling Approach | Correlation | MAPE (%) | Predicted Exceedances (%) | False Alarm Ratio (%) | Mean Density for False Alarms ($\mu\text{g}/\text{m}^3$) |
|---|-------------|----------|---------------------------|-----------------------|--|
| SARIMAX model | 0.91 | 18.62 | 76.97 | 46.45 | 172.14 |
| Artificial Neural Network | 0.91 | 16.53 | 71.56 | 45.48 | 173.14 |
| Esemble: max [Pred _{SARIMAX} , Pred _{ANN}] | 0.92 | 19.32 | 82.15 | 50.91 | 170.38 |

Figures. 11-13 compare the behaviours of observed and simulated concentrations during the validation period in relation to the above three models. The scatter plots on the right-hand side show similar and high levels of linear correlation between observations and forecasts for all models, while the line plots on the left-hand side demonstrate that the artificial neural network has the worst performance in predicting pollution peaks. This can be explained considering that the training of a neural network does not use the entire available dataset to estimate the vector of weights associated to the neural connections; in fact, the test subset of data (containing in this application as many as 1712 observations) is not employed in the learning process, but is used to conduct an unbiased estimation of the network's likely performance.

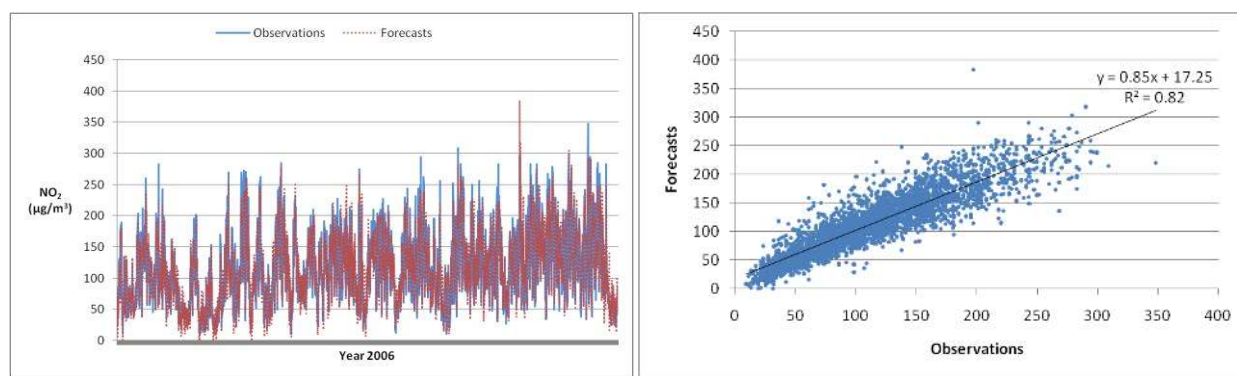


Fig. 11. NO_2 concentrations observed in 2006 at Marylebone road versus SARIMAX forecasts.

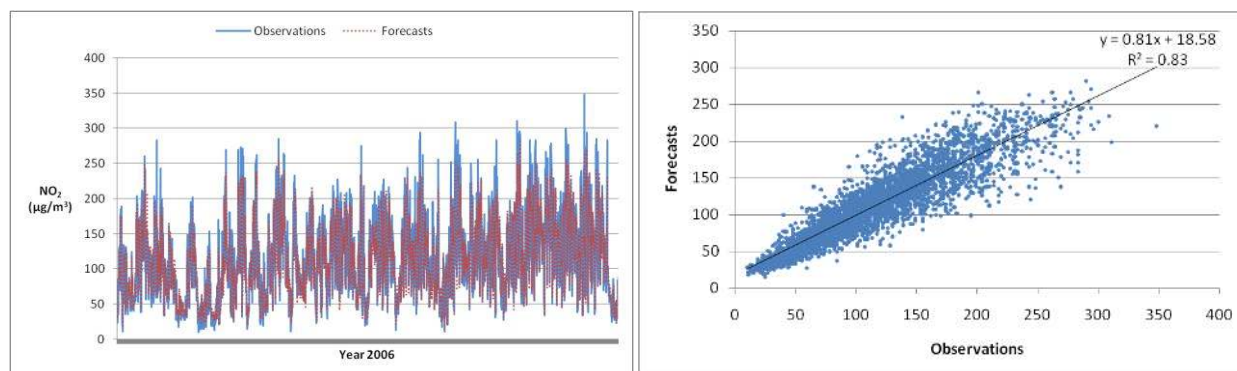


Fig. 12. NO₂ concentrations observed in 2006 at Marylebone road versus Artificial Neural Network forecasts.

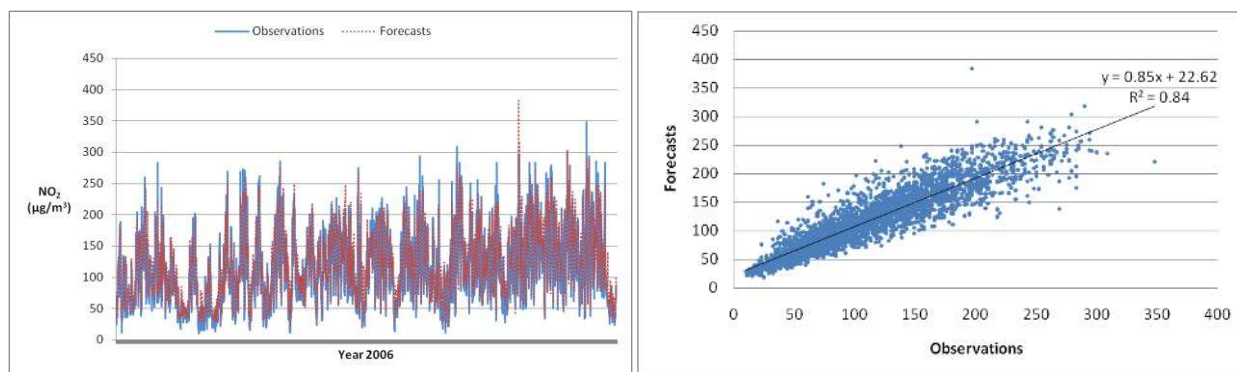


Fig. 13. NO₂ concentrations observed in 2006 at Marylebone road versus Ensemble forecasts.

To check the robustness of the two forecasting models in comparison, a sensitivity analysis has been performed on the influential factors, that are traffic volume, wind speed and wind direction. In detail, as these variables are likely subject to measurement errors, the models' prediction performance has been evaluated under two scenarios: a Low Scenario perturbing the explanatory variables' levels, observed in 2006 at Marylebone road, with a uniform random error that ranges from -5% to +5%; a High Scenario based on a -15% to +15% uniform random error.

Table 8 compares the performance indicators' values under the two scenarios with those based on the original dataset, in relation to an alarm concentration threshold set at 180 µg/m³. It emerges that the assumed perturbations impact the accuracy of both models slightly, which means that these are very robust to measurement errors.

Table 8. Sensitivity analysis results if the alarm threshold is 180 µg/m³.

| Modelling Approach | Correlation | MAPE (%) | Predicted Exceedances (%) | False Alarm Ratio (%) | Mean Density for False Alarms (µg/m ³) |
|---|-------------|----------|---------------------------|-----------------------|--|
| SARIMAX model | 0.91 | 18.62 | 76.97 | 46.45 | 172.14 |
| SARIMAX model-Low Scenario | 0.91 | 18.87 | 76.64 | 46.19 | 172.51 |
| SARIMAX model-High Scenario | 0.90 | 20.17 | 76.32 | 46.54 | 172.85 |
| Artificial Neural Network | 0.91 | 16.53 | 71.56 | 45.48 | 173.14 |
| Artificial Neural Network-Low Scenario | 0.91 | 16.60 | 70.94 | 45.30 | 172.85 |
| Artificial Neural Network-High Scenario | 0.91 | 16.71 | 72.50 | 45.79 | 172.77 |

5 Conclusive remarks and future steps

The illustrated findings reveal that the neural network can represent properly the non-linear relationships between concentration and weather, despite its minor ability at foreseeing exceedances with respect to the alternative models. So, it could play a role as a non-parametric tool for a preliminary analysis of the above non-linear behaviours so as to gain knowledge for improving the specification of the better SARIMAX model by the introduction of non-linear components.

This initial study also demonstrates that the ensemble approach is really promising for the prediction of extrem pollution events. This is surely an area of research that deserves further investigation. In particular, an interesting issue to tackle is the influence of the number and type of ensemble members on the performance in forecasting pollution peaks.

Moreover, we believe that forecasts of high concentrations could be further improved by employing as transport-related causal factor a variable representing the mean hourly behaviour of congestion. The relevant literature, in fact, states that NO₂ concentration is more related to variability in vehicle speed than to speed itself or traffic flow (Tartaglia, 1999). Hence, it is plausible that extreme levels of pollution can be better explained with stop and go driving and tailback formation.

On the side of meteorological influence over NO₂ concentration, we intend to evaluate the explanatory power of some other atmospheric variables, notably solar radiation, concentration of volatile organic compounds and vertical temperature gradient. The first is radiant energy emitted by the sun, particularly electromagnetic energy, and is an influential factor of photochemical smog, along with volatile organic chemicals. In fact, nitrogen dioxide from vehicle exhaust is photolyzed by incoming solar radiation to produce nitrogen oxide and an unpaired oxygen atom, which then combines with an oxygen molecule to produce ozone. Under normal conditions, the majority of ozone particles oxidize nitrogen oxide back into nitrogen dioxide, thus leading to an only temporary increase in ozone density near ground level. However, when volatile organic compounds are present in the air, they turn nitrogen oxide into nitrogen dioxide without breaking down any ozone molecules, which rises ozone levels durably.

The vertical temperature gradient, instead, is important for the formation-dispersion process of all air pollutants. It measures how much colder the air gets as we move vertically up through the atmosphere. If the temperature decreases slowly with altitude, or even increases, the atmosphere becomes stable and slackens the dispersion of pollution.

A final suggestion concerns the possibility of enhancing the ability to foresee exceedances by requiring traffic management actions when the predicted concentration overcomes a limit that is pretty high but lower than the normative one (200 µg/m³). This improvement occurs because a smaller alarm threshold compensates for the tendency of models to underestimate anomalous levels of NO₂ pollution. This additional limit value has been identified with a rule of thumb approach balancing forecasting performance and incidence of false alarms. Further research could investigate the development of an objective function to guide the determination of this other threshold through a mathematical programming method.

Another future step will focus on the prediction of extreme NO₂ pollution events for sites where these phenomena happen rarely, but with a frequency that can result in violation of the air quality standards. In such cases, previous studies (Stockwell et al., 2002) have highlighted the limitation of site-specific statistical models. The hypothesis to be tested is that the application of a panel data dynamic econometric model may outperform a site-specific model at predicting pollution peaks. This assumption is underpinned by the idea that merging the datasets describing the behaviour of NO₂ hourly density in many different places of a Region or a Country can increase the frequency of observed severe pollution events along with the variability of transport and meteorological variables, thus resulting in more accurate forecasts of concentration during periods of unusual emissions and/or weather conditions. Further, the panel framework would permit to introduce the geographic characteristics of the monitoring sites as influential factors (e.g. the ratio of average buildings' height to street width), thus making the prediction model highly transferable.

We also intend to explore the possibility of estimating with fairly accurate precision the yearly frequency distribution of concentrations and related average well in advance of the end of the year using a sample of the annual set of observations (for instance, the data collected during the first few months). This hypothesis should be valid for transport-related air pollutants, as mobility patterns have a cyclical nature and are rather stable. If the discussed assumption will be demonstrated, a novel methodology able to anticipate the annual mean concentration and its distribution for different sites will emerge, enabling Local Authorities and policy makers to timely define and implement medium-term strategies (modal shift plans, low emissions zones, relocation of queues, etc.) and traffic management options to prevent or reduce the exceedance of pollution limits.

In the end, we will investigate the feasibility and the potential benefits of the interaction between real-time air quality forecasting and traffic microscopic simulation, to mitigate atmospheric pollution exposure in line with the environmental legislation requirements. This is based on the idea that, if a concentration peak (that may go beyond the regulatory threshold) could be predicted few hours in advance, solutions to avoid its occurrence should be sought by a quantitative approach. As occurs in the real-world applications of traffic management, data coming from a real-time traffic surveillance system (UTMC, SCOOT, etc.) could be fed into a dynamic microsimulator for short-term estimation of network status in terms of flows, average speed of vehicles, travel times and traffic jams. These traffic-related predictions could be then used to foresee air quality. If an exceedance of any concentration limit were foreseen, the impact of alternative traffic management scenarios on pollution could be simulated through the interplay between the microsimulator and

the air quality forecasting model. This might aid the choice of an ex-ante strategy to prevent the extreme pollution event at the minimum social cost in terms of effects on important transport variables like, for instance, network congestion, accessibility and public transit service quality.

References

- Baur, D., Saisana, M. and Schulze, N. (2004), "Modelling the effects of meteorological variables on ozone concentration-a quantile regression approach". *Atmospheric Environment* 38, pp. 4689-4699.
- Bell, M., Bergantino, A.S., Catalano, M., Galatioto, F. and Migliore, M. (2015), *Previsione dell'inquinamento generato dalla mobilità veicolare e traffic management*, *Rivista di Economia e Politica dei Trasporti*, *forthcoming*.
- Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Oxford University Press.
- Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. 2008. *Time Series Analysis: Forecasting and Control*. 4th ed. Hoboken, NJ: Wiley.
- Cai, M., Yin, Y., Xie, M. (2009), "Prediction of hourly air pollutant concentrations near urban arterials using artificial neural network approach", *Transportation Research Part D* 14, pp. 32-41.
- EU Press Release Database (2014), "Environment: Commission takes action against UK for persistent air pollution problems", <http://europa.eu/rapid/press-release_IP-14-154_en.htm>.
- European Union (2008), Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe.
- European Environment Agency of the European Union (2014), *Air Quality in Europe - 2014 Report*, Publications Office of the European Union, Luxembourg.
- Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.
- Heinrich, J. et al. (2005), "Studies on health effects of transport-related air pollution", in Krzyzanowski, M., Kuna-Dibbert, B., Schneider, J (eds.) *Health Effects of Transport-Related Air Pollution*, World Health Organization-Europe, Denmark.
- Hunt P.B. et al. (1981), *SCOOT - A Traffic Responsive Method of Coordinating Signals TRRL report LR1014*, Crowthorne.
- Inderjeet Kaushik and Rinki Melwani (2007), "Time series analysis of ambient air quality at its intersection in delhi (india)", *Journal of Environmental Research and Development* 2, pp. 268-272.
- Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., Niska, H., Dorling, S., Chatterton, T., Foxall, R. and Cawley, G. (2003), "Extensive Evaluation of Neural Network Models for the Prediction of NO₂ and PM₁₀ Concentrations, Compared with a Deterministic Modelling System and Measurements in Central Helsinki". *Atmospheric Environment* 37, pp. 4539-4550.
- Lavecchia, C., Pilati, S., Angelino, E., Fossati, G. (2007), "Analisi dei dati di traffico esistenti per la definizione dei profili temporali: metodologia ed esempio di applicazione", XIII Expert Panel Emissioni da trasporto su strada, Roma, 4 ottobre 2007.
- Nagendra, S.M. S., Khare, M. (2006), "Artificial neural network approach for modelling nitrogen dioxide dispersion from vehicular exhaust emissions", *Ecological Modelling* 190, pp 99-115.
- Perez, P., Trier, A. (2001), "Prediction of NO and NO₂ concentrations near a street with heavy traffic in Santiago, Chile", *Atmospheric Environment* 35, pp. 1783-1789.
- Re, M., Valentini, G. (2012) "Ensemble methods: a review". In: *Advances in Machine Learning and Data Mining for Astronomy*, Chapman & Hall Data Mining and Knowledge Discovery Series, Chap. 26, pp. 563-594.
- Sayegh, A. S., Munir, S., Habeebullah, T. M. (2014), "Comparing the performance of statistical models for predicting PM₁₀ concentrations", *Aerosol and Air Quality Research* 14, pp. 653-665.
- Seinfeld, J.H., Pandis, S.N. (2006), *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, John Wiley & Sons, Inc., ISBN 978-0-471-72018-8.
- Shumway, R.H. (1988), *Applied Statistical Time Series Analysis*, Englewood Cliffs: Prentice Hall.
- Stockwell, W.R., et al. (2002), "The scientific basis of NOAA's air quality forecast program", *Environmental Manager* 8, pp. 20-27.
- Tartaglia, M. (1999), *L'Inquinamento dell'Aria da Traffico Stradale*, Ed. Bios.
- The In-House Lawyer (2014), "Failing health: why air quality legislation is not working", <<http://www.inhouselawyer.co.uk/index.php/environment/10524-failing-health-why-air-quality-legislation-is-not-working>>.
- Verbeek, M. (2004), *A Guide to Modern Econometrics*, John Wiley & Sons Ltd, England.
- Viotti, P., Liuti, G., Genova, P.D. (2002), "Atmospheric urban pollution: applications of an artificial neural network (ANN) to the city of Perugia", *Ecological Modelling* 148, pp. 27-46.
- VV. AA. (2008), Special Issue on Applications of Ensemble Methods, *Information Fusion Journal* 9 (1).
- World Health Organization (2014), *Ambient (outdoor) air pollution in cities database 2014*, <http://www.who.int/phe/health_topics/outdoorair/databases/cities/en/>.
- Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., Baklanov, A. (2012), "Real-time air quality forecasting, part I: History, techniques, and current status", *Atmospheric Environment* 60, pp. 632-655.